Philosophical discussions of responsibility often begin with a contrast between two ways in which the phrase “X is responsible for Y” gets used. On the one hand, the initiator of the discussion might say, there is a purely causal use of the term: “X is responsible for Y” in this sense picks out X as the cause, or at any rate as a salient cause, of Y. Anything that can be a cause can therefore stand in the place of Y – a person, an animal, a piece of machinery, a religious belief, a weather phenomenon. On the other hand, there is what is typically described as a moral use of the term: “X is responsible for Y,” in the moral sense, asserts a tighter or deeper connection between X and Y, of a kind that would make blaming X appropriate, if the Y in question is bad, and praising X appropriate if the Y is good. The kind of being that can occupy the X place in the moral sense may be called a morally responsible agent. It is generally agreed that inanimate objects, lower animals, young children, not to mention ideas and weather phenomena, can be causally but not morally responsible. If there are any morally responsible agents at all, people think, the best candidates are adult human beings of average or better than average intelligence and mental health.

I trust that this contrast will be familiar to most of you, or, if not familiar, at least easily grasped. To reinforce our sense of the contrast, consider a state of affairs – let it be an undesirable state of affairs – that could have any of a variety of possible causes. A stain on the carpet, for example, could be the result of a pipe’s bursting, a baby spitting up, a cat knocking over the milk, or a roommate carelessly spilling coffee on her way to her desk. In the causal sense, it could be a pipe or a baby, a cat or a roommate that is responsible for the mark. But it is only the last, it seems natural to say, that it would be appropriate to blame for the spot, for only the roommate is or could be morally responsible for the stain.
Still, sticklers for language, as philosophers tend to be, might point out that my characterization of the contrast is too rough. For just as there is a broader and a narrower sense of *responsibility*, there is a broader and a narrower sense of blame. It would not be unusual for someone to “blame” the weather for spoiling her picnic, or the carburetor for making the car stall. Similarly, it would not be weird or confusing for someone to blame the stain on the carpet on the cat, even though in a stronger or deeper sense, only the roommate would be an appropriate object of blame. And indeed, these two senses of blame seem exactly parallel to the two senses of responsibility. When we blame an event on a cat or a carburetor, we mean only to say that the individual is causally responsible for the event, under the condition, implied by our choice of the word “blame,” that we judge the event in question to be undesirable or bad. And of course we *can* use this same causal sense when we blame an event on a person – when, for example, a host’s sudden illness requires that the party be cancelled, or one’s enjoyment of a concert is spoiled by a coughing fit of someone sitting nearby. But when someone is morally responsible for an event that is bad or wrong, a different kind of blame may be in order, which we may call moral blame.

For the remainder of this paper I shall use “blame” and “responsibility” to refer to the narrower, stronger, or “moral” sense of these terms. It is the concepts to which these stronger senses refer that have been and continue to be the subject of so much philosophical puzzlement and debate. But I opened my discussion with the contrast between the broader and narrower uses of the terms not only to clear the ground but also to bring out what I take to be a pervasive assumption in discussions of responsibility and blame – namely that there is a dichotomy, or at any rate, a single, unitary contrast in the background between responsible and nonresponsible
beings, between those individuals who can be appropriate objects of blame and those who cannot. ¹

For much of the philosophical topic’s history, the contrast was characterized in terms of whether individuals who were causally responsible for things were appropriately or justifiably deserving of punishment, reward, condemnation, and approval. But in the 1960’s, the philosopher P.F. Strawson significantly enlarged and enriched the discussion by emphasizing the attitudes beside and beneath these responses that often animate our interest in a deeper way. As Strawson noted, the “practices or attitudes [just mentioned, of punishment, condemnation, reward, and approval] permit, where they do not imply, a certain detachment from the actions or agents which are their objects.” (Watson 75) Strawson turned our focus instead toward “the non-detached attitudes and reactions of people directly involved in transactions with each other; [he wrote] of the attitudes and reactions of offended parties and beneficiaries; of such things as gratitude, resentment, forgiveness, love, and hurt feelings.” These attitudes “of involvement or participation in human relationships” Strawson termed “reactive attitudes.” He contrasts them with what he calls the objective attitude (or range of attitudes),” which he characterized as follows:

To adopt the objective attitude to another human being is to see him, perhaps, as an object of social policy; as a subject for what, in a wide range of sense, might be called treatment; as something certainly to be taken account, perhaps precautionary account, of; to be managed or handled or cured or trained; perhaps simply to be avoided… The objective attitude may be emotionally toned in many ways, but not in all ways; it may include repulsion or fear, it may include pity or even love, though not all kinds of love…it cannot include the range of reactive feelings and attitudes which belong to involvement or participation with others in inter-personal human relationships; it cannot include resentment, gratitude, forgiveness, anger, or the sort of love which two adults can

¹ David Shoemaker is a notable exception to the philosophers whose unitary view of responsibility I am questioning in this paper, and he has defended ideas similar in many respects to the ones I put forward here. In a more polished draft, which will include more footnotes than this one (and not this one), I will have to say more about the relation between my views and Shoemaker’s. Apologies to David for not having done so in time for the discussion on December 3.
sometimes be said to feel reciprocally, for each other. If your attitude towards someone is wholly objective, then though you may fight with him, you cannot quarrel with him, and though you may talk to him, even negotiate with him, you cannot reason with him. You can at most pretend to quarrel, or to reason, with him. (79)

Part of Strawson’s point is to remind us how central it is to our form of life that we engage with each other reactively, to point out how much colder and bleaker our lives would be if we took only the objective attitude (or range of attitudes) toward ourselves and each other instead. Another part is to make the connection between our proneness to see ourselves and each other as subjects and objects of the reactive attitudes, on the one hand, with the concepts of responsibility and blame on the other. Indeed, to many philosophers (including myself) who came of intellectual age since the publication of Strawson’s essay, it has seemed reasonable to identify being a responsible agent with being an appropriate object of reactive attitudes, and to identify blameworthiness, more specifically, with being an appropriate object of the negative reactive attitudes of resentment, indignation, anger, and the like.

Strawson’s essay, while it deepens our understanding of responsibility and blame, accepts the idea that there is a single contrast in the background: between two types of agent - the responsible and the nonresponsible, two types of relation an agent may bear toward her actions and their consequences, and two types of attitude – reactive attitudes and objective attitudes – that it may be respectively appropriate to have. Strawson recognizes that the categories that he contrasts do not admit of sharp boundaries. There is, as he notes, a continuum stretching between individuals, such as inanimate objects and infants, at one end and mature, mentally healthy adult human beings on the other. Still, his writing suggests that there are two clusters of individuals, or agents, and a contrast between them, even if there is a continuum running from one pole of the contrast to the other.
If an individual is responsible it can be appropriate to punish and reward, to blame and credit, and to hold reactive attitudes towards her, and if an individual is not responsible then none of these reactions are justified. Although there is enormous controversy about how to understand the concepts on the one side of the contrast, on whether these concepts properly apply to anyone, and if so, when and to whom, it is generally assumed that the concepts over which such controversies are discussed are themselves single, unified concepts that all parties to the discussion are trying to analyze and understand. To ask what responsibility is, what blame is, what the category of reactive attitudes is, is to assume that the term or phrase in each case refers to some single feature, phenomenon or category, if it refers to anything at all. My aim today is to question this assumption and to do so by focusing on the reactive attitudes – more specifically, on the negative reactive attitudes that are commonly regarded as forms of blame. I shall be arguing that there are at least two sets of negative attitudes that we sometimes have toward individuals, both of which are to be contrasted with objective attitudes, but which have different conditions of appropriateness. We are used to thinking that if an individual is responsible it may be appropriate to punish and reward, to blame and credit, and to hold reactive attitudes towards her, and that if an individual is not responsible then none of these reactions are justified. But if the reactive attitudes do not constitute a unified category, with a unified set of conditions, then it becomes unclear what we so much as mean by blame and credit, what we mean by the idea of responsible agency, and how these concepts relate to the appropriateness of punishment and reward.

So let us turn our attention to blame and to the attitudes that are thought to constitute forms of blame. What is it to blame someone? We may begin by considering some paradigm cases: A woman is late picking up her friend to go to the movies, causing them to miss the
previews and even the opening scene. The friend is annoyed, and thinks the driver should apologize. This case is trivial, but there are equally clear cases at the more serious end of the spectrum of importance. A man discovers that his nephew is poisoning his father, wanting to kill him and secure his inheritance. The man is furious and indignant, and wants the nephew to be severely punished. In both cases, two central elements in the blamer’s perception of the situation combine to lead to a judgment of blame: (1) An individual does something wrong or harmful that is (2) due to a fault in the individual’s character. Together, these lead the blamer to form a negative attitude (annoyance and anger respectively) and a desire or demand for some sort of sanction (in the first case, that the friend apologize, in the second, perhaps jail time).

But these two elements might also occur independently. An individual might do something wrong or bad, though no fault in the individual is involved, or an individual might possess a fault that yet fails to issue in a wrong or bad action. In such cases, is it still possible or warranted to blame? One might think that the answer in the first case is easy: If the action does not issue from an individual’s fault, blaming the individual anyway would be wrong or mistaken. In fact, I believe that the issue is more complicated but I shall not pursue this here. Instead, let us turn to the other side of the question: If a person has a fault, but does nothing wrong – or at any rate nothing one knows about that is wrong or harmful to oneself or to people one cares about – might one justifiably blame the person anyway? Imagine for example that a woman knows that her friend is the sort of person who is perennially late. She shows up at her classes at the very last minute; she’s the last to arrive at a party, and so preemptively the woman always volunteers to drive when they go to the movies together. Or consider the nephew: He is, let us say, a nasty piece of work – selfish, greedy, and callously indifferent to the fate of his father. If someone proposed a way for him to murder his father that had a good chance of going
undetected, he would agree to it in a minute, but he is not very bright or imaginative and so no plan or attempt materializes. When the woman reflects critically on her friend’s lack of punctuality or the uncle on his nephew’s vicious character, do their judgments and attitudes constitute instances of blame even though there is no particular event or action they are blaming the respective objects of their reflections for?

Opinion might divide in answer to these questions, and many of us might feel at least initially torn. On the one hand, it seems linguistically required that when one blames someone there must always be some act or state of affairs that one is blaming the individual for. And it seems possible simply to notice that someone has a fault, or even a vice, without necessarily blaming them for it. One might observe that a person has a certain trait, which one registers as undesirable or bad, and just leave it at that - one may take notice of a person’s flaws, for instance, but regard them with Strawson’s ‘objective attitude.’

On the other hand, we are not always so dispassionate and sanguine when we consider people’s faults, especially when the people in question are family members or friends. Often we do not just notice that someone has a fault – we criticize him for it. That is, we take a critical attitude toward him, which we may or may not express. Criticism, like blame, seems to involve or at least license one’s taking a negative affective attitude of some kind. One might be annoyed by a colleague’s laziness; disgusted by a friend’s sloppiness; appalled by a sibling’s political opinions. And people do sometimes speak of blaming a person simply for being a certain way. Indeed, some philosophers suggest that blame is always at bottom of this sort. Though we might say we are blaming someone for acting selfishly or obnoxiously, they would argue, the act is merely the occasion which calls attention to the fault for which the person is really being blamed.
Is criticism, then, a species of blame or is it something different? I shall argue that it is
different, but, lest one think that this is merely a question of linguistic niceties, let me explain
why I think there is something philosophically significant at stake.

Earlier I drew attention to a cluster of ideas that we commonly employ to contrast two
kinds of individuals and two corresponding sets of attitudes and judgments that we apply toward
them. Some individuals are responsible agents; others are not. The appropriateness of blame
and praise, of punishment and reward, of the reactive attitudes, are restricted to the domain of
responsible agents. Nonresponsible individuals, by contrast, may be proper objects of social
policy; they may be ‘managed or handled or cured or trained,’ they may be pitied or feared,
avoided or enjoyed, but only in ways that can be embraced while taking an objective attitude
toward them. If we accept the contrast, then the question of how to classify or understand
criticism presents us with two options.

If we think of criticism on the model of one who simply, dispassionately notices that an
individual has a regrettable trait or feature – we can call this the “just sayin’” model, then we
may place criticism on the non-responsible side of the divide. To criticize, according to this
conception, is to ascribe a negatively valued feature to someone. When one criticizes something,
one need make no assumption that the object of criticism is a responsible agent, much less that
the individual is responsible for the specific trait being ascribed to it. One can, on this
conception, criticize a book for being boring, or a train for being slow; further, one can criticize a
person for being clumsy or stupid or ugly, as well as for being selfish or pompous or mean.

On the other hand, if we think of criticism as involving something more than mere
negative description, and as involving or licensing affective attitudes like annoyance, (moral)
disgust and contempt, then it seems we should place it on the opposite side of the contrast.
Criticism, on this conception, is a kind of blame, or at least a kind of response to an object that is in the neighborhood of blame. The attitudes that seem reasonably to accompany or follow from the attribution of viciousness to an individual are, according to this view, as reactive as those that respond to specific acts that reflect such viciousness. To regard someone as an appropriate object of criticism thus presupposes that the individual is a responsible agent just as much as does regarding someone as an appropriate object of paradigmatic blame.

But if, as I think, neither of these conceptions of criticism is acceptable, it suggests that there is something wrong with the contrast that dictates that these are the only options there are. There is, I shall argue, a quite familiar kind of criticism that cannot be accommodated by the Just Sayin’ Model. It is not dispassionate, and it is in tension with the objective attitude with which paradigmatically nonresponsible beings must be regarded. But neither does it conform to the standard model of blame, and the differences between it and standard blame are not trivial. Importantly, the conditions that justify standard blame are different from those that make criticism appropriate – so different that one can conceive of beings who are capable of deserving the one but not the other and vice versa. This suggests that the concept of responsibility is not so unified and straightforward as it is typically assumed to be, and that there is not one contrast but two between the objective attitude and the attitudes of involvement or participation in human relationships. In light of this, as we’ll see, new puzzles arise about how we should understand punishment, as well as about how to think about the age-old philosophical problem of free will.

Just as there are broader and narrower uses of the terms “responsibility” and “blame,” there are broader and narrower uses of “criticism” that make it difficult to clearly establish the features of the category I wish to call attention to. Still, I hope that the inadequacy of the Just Sayin Model of Criticism as a category that captures all standard uses of the term will be
relatively easy to see. When we observe that a person has bad skin, for example, we attribute a fault or defect to her which we nonetheless don’t criticize her for. If we believe that Kant was a brilliant philosopher, but that his metaphysical and ethical views are nonetheless mistaken, we might criticize Kant’s philosophical positions without thereby criticizing him. And although it seems reasonable and appropriate to criticize a person for being indifferent to the suffering of others, it seems totally misguided to criticize a crocodile or a killer whale for this trait. But since bad skin, bad metaphysics, and indifference to human suffering are undesirable features for an individual to have, this shows that criticism, in one sense at least, must involve something more than the attribution to an individual of an undesirable trait.

These considerations, indeed, all seem to support the idea that criticism (or at least an important sense of criticism) belongs on the other side of the contrast, to be classified, if not as a form of blame, at least as being in the neighborhood or family of blame. Like blame, criticism involves or allows affective attitudes like annoyance and contempt; like blame, only certain sorts of individuals (nephews, yes; crocodiles, no) seem capable of deserving it. It seems plausible to think that regarding someone as an appropriate object of criticism presupposes the individual’s being a responsible agent just as much as does regarding him as an appropriate object of blame. And our criticisms of ourselves and each other, for being, for example, closed-minded, pretentious, obnoxious, or cruel, along with our pride and admiration for people who are generous, kind, imaginative, or brave, fit just as much into “the range of reactive feelings and attitudes which belong to involvement or participation with others in inter-personal human relationships” as do the reactive attitudes Strawson explicitly mentioned such as resentment, indignation, and gratitude.
For these reasons, I am inclined to classify criticism and shame, admiration and pride, or at least the kinds of criticism, admiration, and so on with which I am concerned, as belonging to the range of reactive attitudes that reflect our interaction with others as potential members of our interpersonal communities. To regard someone as a fit object of criticism, then, is to regard him as being an individual of a certain kind – which for the moment we can identify with responsible agency. And to regard someone as a fit object of criticism for having a particular feature or for committing a particular act, is, correspondingly, to regard him as having a kind of relationship to that feature or act which, for the moment, we can call one of responsibility (for that feature or act). This seems to explain or reflect the fact that it is appropriate to criticize the nephew but not the crocodile for his indifference to human suffering as well as the fact that it is appropriate to criticize one’s friend for always being late, but not for always sneezing during hayfever season. But before agreeing that these considerations place criticism in the neighborhood or family of blame, we need to look more closely at what it is about the difference between a person and a crocodile, and about a person’s relation to tardiness and her relation to hayfever, that makes them respectively criticizable or not.

In virtue of what is it appropriate or fitting that we can criticize a person for being indifferent to suffering, but not appropriate to criticize a crocodile or killer whale? And in virtue of what is it appropriate to criticize a person for being perpetually tardy, but not appropriate to criticize a person for being allergic to pollen? It is tempting to answer such questions by reference to people’s alleged control over the traits and acts for which they are being considered for criticism. Crocodiles can’t help the fact that they are indifferent to suffering, after all, and it is not up to us whether we have allergies. And so we might be inclined to think that this is the
key to explaining the difference in our reactions to these cases and those with which we contrast them.

But I think the answer has less to do with the individual’s control over what he is or does than with the kinds of reasons for being or for acting differently to which he is susceptible. Thus, for example, at the risk of being frivolous, let us first reconsider the crocodile. It may be true that the crocodile can’t help being indifferent to human suffering, but what if he could help it? The crocodile, I would argue, has no reason not to be indifferent, and that is sufficient to eliminate any basis for criticism independently of the issue of control. (Similarly, I would argue that an alien from outer space might have no reason not to be indifferent to us, making it similarly inappropriate to criticize him even though he has as much control over his attitudes and behavior as we do.)

Further, it seems to me that there are cases where criticism is appropriate even when control is lacking. To see this, let’s consider our attitude to a friend who, instead of being perpetually tardy, is an inveterate sexist. He applies double standards in his judgments of men’s and women’s characters; he is much less likely to take women’s ideas seriously than men’s; and he is frequently amused by incidents that one finds appalling. Is one critical of one’s friend for this? Of course. And yet it doesn’t seem to me that we need to assume or believe that our friend’s views about women are subject to his control. He may, for example, have been brought up in a community where these attitudes were nearly universally accepted, shaping his continuing perception in ways he may not even be aware of. If I am right about this, then the salient feature that makes one’s friend subject to criticism is not that it is up to him to be a sexist or not to be; rather, it is that there are reasons not to be a sexist that he has the capacity to appreciate that he is failing to recognize.
The same, of course, could be said of one’s tardy friend –like the sexist, she has reasons not to be tardy that she fails to appreciate, whereas in the case of her allergies the idea that they reflect a failure to appreciate reasons can get no purchase.

This suggestion is just a first stab at characterizing the conditions or assumptions under which criticism may be appropriate. No doubt it needs clarification, explanation, and qualification in order to be fully adequate, but if it is on the right track it suggests a feature of responsible agency that is independently plausible. Specifically, to be a fit subject of criticism, according to this suggestion, one must be an individual that can be sensitive to reasons of a certain sort. In order to be appropriately criticizable for being a sexist or for being cruel, one must be the kind of being that is capable of understanding and appreciating reasons not to be sexist or cruel. Indeed, I am inclined to say that one cannot even be sexist or cruel if one is not that kind of being. (A dog that, for some reason, had been traumatized or programmed to respond differently to men and to women would still not be a sexist; and though crocodiles are indifferent to human suffering, they cannot be callously indifferent.) And, insofar as we identify being a responsible being with being a potential member of an interpersonal community, it is plausible that this should be a condition. For an interpersonal community is a community within which one may find one’s friends and lovers, and it is plausible that such relations require a shared universe of possible perceptions and reasons within which one can communicate and exchange ideas, values, and other responses to the world around one.

This may sound at first as if it is the same condition as that which is needed to make blame appropriate, for it is generally acknowledged that it would be wrong to blame someone for doing something if the person blamed was not able to understand and appreciate why she shouldn’t have done it. But recalling the remarks I made a moment ago about the relation
between criticism and control suggests that the conditions of criticism are different from the
conditions of blame.

My description of the sexist was intended specifically to show that a person could be
appropriately criticized even if he lacked control over the feature he was criticized for. It need
not, at least in the short run, be in the man’s power not to be sexist; this is not a feature that is
under the direct control of his will. And given his upbringing and the make-up of his social
world, we may grant that one could not have reasonably expected him to be (or to have been)
any different. This does not, to my mind, make criticism inappropriate. It is still bad or wrong
to hold sexist attitudes, and the man in question is (by stipulation) still capable of appreciating
the reasons why. Nor need we register his sexism in an exclusively cool and detached way,
taking the objective attitude to, as it were, ‘men of a certain age’ or from a certain background.
We may argue with him, if we think it worth our while – or, to invoke a contrast of Strawson’s,
we may reason with him, and not just negotiate with him. Alternatively, we may recoil from
him, as we recoil from people with other vices, for his being, let us say, a smug, chauvinist jerk.

But although the consideration that his sexism is not under his control does not block our
criticism, it does seem to me to block the appropriateness of blame. Indeed, it is precisely
considerations like these that are likely to make those of us who are initially tempted to blame
someone for his sexism or his sexist behavior, withdraw our blame. Why is this? Many
philosophers have suggested that it is because of the close relation between blame and the
imposition of sanctions. Blaming people licenses demands for such things as apology,
compensation, and punishment, according to the case at issue. And it seems unfair to impose or
demand such things of an individual if she is unable to avoid being subject to them.
This idea is often discussed in connection with the notion of accountability. The philosopher Gary Watson, for example, has suggested that when you blame someone, you hold that person accountable, which is to be understood in terms of practices that assign specifiable tasks or expectations to particular individuals, with an understanding, common to all parties involved, that if the individuals fail to perform these tasks or live up to these expectations, they will be subject to sanctions. Such practices may be more or less explicit. But all cases of accountability refer at least implicitly to a norm or standard, which one party is expected to abide by, and which another party is in a position to demand that she do. Although not every case of holding accountable for something bad need be a case of blaming, every case of blaming is a case of holding accountable.

This helps explain why, in many instances, criticism is appropriate but blame is not, and why if blame is appropriate, it may indicate that a particular kind of relationship is in place. For example, we might criticize someone, but not blame them, for being a slob or for being pretentious or for being a poor philosopher, because even if we disapprove of these traits and the patterns of behavior that issue from them, it may be none of our business whether the object of our criticism possesses them. They are not accountable to us to be neat or down-to-earth or to marshal better philosophical arguments – unless, that is, the slob is your housemate or the bad philosopher an academic opponent spreading a caricaturized version of your views. Further, it can explain why we are less likely to blame someone for being a certain way – for being selfish or dishonest or sexist, for example – than for doing things that manifest these traits – such as stealing or lying or discriminating against women. For we often think that a person can’t help being, as it may be, selfish or mean or sexist, but that they can help acting on their selfish, mean, or sexist impulses, and so blaming them for these actions may not be unfair.
These reflections confirm and shed light on the observation that criticism is distinct from blame, but they do little to support the claims of disunity I made earlier about the concept of responsibility and the category of the reactive attitudes. For it is compatible with the observations I have made so far that one think that blame is distinct from criticism only in requiring a stronger kind of responsibility, and that the conditions of appropriate blame include the conditions of criticism, differing from them only in requiring something more. Indeed, among the relatively few philosophers who agree that there is more than one kind of deep and more than merely causal responsibility, this seems to be the standard position. Gary Watson, for example, has proposed that we recognize two senses of responsibility, which he associates with the notions of attributability and accountability respectively. To regard someone as responsible for something in the attributability sense, he suggests, is to regard the thing as disclosing something about the person’s self. Judgements of personal character, including critical judgments, are examples of attributability. But to hold someone responsible in a way that legitimates blame and sanctions is, as we have already seen, to invoke the idea of accountability. According to Watson, it appears that an individual is accountable for something only if, in addition to its being attributable to him, there is a person or group to whom he is accountable, and the individual has the ability to avoid being subject to the sanctions that failure to conform to the relevant standards or expectations would in these circumstances expose him to.

On such a view, responsibility is unified after all, even though there are weaker and stronger forms of it. Critical judgments and attitudes and their positive correlates (such as admiration and esteem) are appropriate as long as an act or a trait is attributable to someone. When, in addition, the individual has the ability to do otherwise and there are practices of accountability in place, the individual may also be justly subject to blame or credit, punishment
or reward. Being a potential object of criticism, then, is conceived as a necessary but insufficient condition of being accountable and blameworthy. And, if one thinks that considerations of determinism imply that no one is responsible in the strong sense necessary for accountability and the justifiability of blame, it may be some consolation that people can at least be seen to possess a weaker kind of responsibility sufficient to justify judgments and attitudes of criticism and admiration.

But this view is mistaken. Once we have teased apart the conditions of criticism from the conditions of blame, and more specifically, once we have seen how closely the appropriateness of blame is linked to accountability, we can see that the category of individuals who can be accountable for things and that of beings who are subject to criticism do not line up as neatly as this view would claim. Moreover, we find reason to question the idea that blame is a stronger attitude, or range of attitudes, than criticism, resting on a stronger set of conditions, so that even if blame is rendered problematic by the problem of determinism, criticism and its opposite are not under threat.

To persuade you, let us look at some kinds of individuals who are appropriately held accountable but who are not paradigmatic of fully responsible agents.

Recall that for individuals to be accountable there must be practices in place that hold individuals to sanctions if they do not conform to specifiable norms or expectations. To be an appropriate candidate for accountability requires the ability to understand these practices and to control one’s behavior in light of them. Normal adult human beings, therefore, are paradigmatic candidates for accountability. But now consider organizations, such as corporations and states. Organizations like these are capable of registering that they are expected to conform to specified norms and standards; they are capable of recognizing that they will be subject to sanctions and
other costs if they do not comply. Further, they typically have decision procedures in place that allow them to choose whether to abide by the norms and standards expected of them or not, in light of the costs they risk incurring if they violate these norms. And indeed, in practice, we do hold corporations and states both morally and legally accountable for things, demanding that in certain circumstances, they compensate victims of their objectionable actions and policies, punishing them with sanctions, boycotts and public condemnations, and even in some cases demanding that they apologize. So, if being responsible is to be identified with being such as to be potentially held accountable for things, corporations and states are responsible. Further, we seem to regard corporations and states as appropriate objects of at least some kinds of blame. In addition, we might consider the status of sociopaths, understanding sociopaths to be individuals completely lacking the capacity for empathy. Like organizations, sociopaths can participate in practices of accountability: they can understand what norms they are expected to conform to and what penalties they are liable to if they do not comply, and in light of this, they can weigh the costs and benefits of following these norms in deciding what to do. If they violate these norms and get caught for doing so, it is not unfair to hold them accountable – they knew what they were risking. But it seems to me that reactive attitudes toward them on the basis of these violations would be completely out of place. When we resent or feel indignant or get angry at people for ignoring or even intentionally defeating our interests, it is because we think they are failing to attend to reasons they have for taking us into account, reasons perhaps for recognizing and respecting us as ends in ourselves (having to do with our status as the offender’s moral equals). People who lack the capacity for empathy however, have no more reason to regard us this way than crocodiles. We should regard such people – as we regard crocodiles – with only the objective attitude. What needs to be emphasized is that whereas they are like crocodiles in being
appropriately regarded with only the objective attitude, they are quite *unlike* crocodiles in being capable of being, and being held to be, accountable. Our concept of responsibility thus seems to break apart – for, if we identify responsibility with accountability, sociopaths will be classified as responsible beings; but insofar as responsibility is associated with being an appropriate object of reactive attitudes, sociopaths are excluded.

What about the converse possibility – are there individuals who can be appropriate objects of reactive attitudes but who are not capable of being accountable? We have already seen, as in the case of our inveterate sexist, that a person may not be accountable for a particular feature that can nonetheless be a basis for the reactive attitudes that accompany criticism. One may be morally disgusted by a person’s sexism without thinking he is appropriately subject to blame for it. But although the sexist is not accountable for his sexism, we may presume that he is accountable for much else (including abiding by laws and policies prohibiting sex discrimination and sexual harassment). Still, I am skeptical that the abilities needed for accountability are *always* necessary for criticizability and attributability. Among the sorts of cases that support my doubts are young children, people with certain sorts of cognitive disabilities, and those with (certain kinds of) mental illness. For people in any of these categories may lack a degree of capacity for self-control or for practical deliberation necessary to be fairly held accountable for acting in accordance with specifiable norms. And yet this does not keep them from having characters that reflect their sensitivity and appreciation of reasons to, for example, share one’s toys with one’s sister or to try to help a stranger who has fallen in the street. The character assessments that Watson discusses that can ground judgments and attitudes of criticism and disdain and of praise and affection, - judgments and attitudes that are as important to being able to include someone in one’s interpersonal community as the ability to
hold them accountable - may be less closely connected to the capacities of practical deliberative reason and control than one might have supposed.

To sum up then, it seems to me that there is at least one significant fault line in the idea of responsibility that mostly goes unrecognized. One important cluster of capacities make possible qualities that ground judgments and attitudes of a certain warm, reactive kind of criticism and praise. These are capacities, as I put it, to be sensitive to and appreciative of (and indeed to have) certain sorts of reasons, capacities that distinguish most adult human beings from infants and crocodiles, but also from sociopaths and corporations. There is, I believe, a strong case to be made for the idea that to be a deeply responsible being is to be the kind of being who is appropriately criticizable. Although I do not have the time to argue for this today, I believe that only beings of this sort are capable of having a genuine character, and of having the kind of self whose disclosure Watson’s notion of attributability-responsibility is meant to capture.

Individuals who lack these capacities entirely are to be viewed with only the objective attitude – they can be “objects of social policy”, they can be “managed and handled…or perhaps simply avoided,” but they are not rightly objects “of the range of reactive feelings and attitudes which belong to involvement or participation with others in inter-personal human relationships."

But curiously (and somewhat surprisingly) some of the individuals in this latter group can be held accountable for things. We may expect them to act according to certain norms, and sanction them if they fail to do so. This suggests that the capacities that make one a potential member of practices of accountability constitute a second cluster. This cluster, too, distinguishes most adult human beings from infants and crocodiles, but for different reasons. And although it does not exclude sociopaths and corporations, it may exclude three-year olds and schizophrenics, while the other group may not.
Where then shall we locate the appropriateness of blame? Is blame a form or expression of accountability? Is it a kind of criticism? Or is blame a combination of both practices, dependent on having the capacities necessary for both accountability and criticism? The language of blame is elastic enough that a case can be made for any of these answers.

It is too late in the day to examine and weigh the reasons supporting one choice rather than another. But because I mentioned near the beginning of this paper that I thought the ideas that I would be presenting had implications for the way we understand punishment and for the problem of free will, let me conclude with some very brief remarks to indicate what I meant by that.

Regarding punishment, my main thought was this: Since punishments are sanctions, theories and practices of punishment must be concerned chiefly with issues of accountability: If an individual is accountable, he is potentially a fit subject of punishment, and his punishment should be set by the specifics of the practices of accountability that are in place. How good or bad his character is, in virtue of which he may have been led to commit an act deserving of punishment is not directly relevant to the question of whether he is accountable and for what. But the hotter, reactive attitudes that typically inform our desire for punishment, and our judgment of how much punishment an offending agent deserves typically (if not always) depend on our criticism of the offender. To fairly determine whether and how much an individual is deserving of punishment, we should distinguish our assessments of whether the individual is accountable for his behavior from that of whether and how much he deserves criticism for it, and we must ask whether it is the business of the punishing agent (a parent, an employer, a state) to criticize as well as to punish. More generally, we may need to develop an ethics of criticism, distinct from an ethics of punishment, that looks not only at the conditions under which criticism
may be deserved but also at the ethical principles and concerns that should govern the public expression of criticism.

And what about the problem of free will? For centuries that problem has been centrally occupied with the question of whether the possibility of determinism (or alternatively mechanism) compromises our ability to do otherwise, or undermines the belief that what we do is under our control in a relevant sense. These concerns are taken to raise questions about whether people can ever be deserving of blame and punishment – whether, in light of the distinction I have been arguing for, they can ever rightly be held accountable for anything. But once we have distinguished accountability from criticizability, this way of understanding the problem seems puzzling. Earlier I pointed out that the kinds of individuals that can participate in practices of accountability include not only adult human beings, but organizations like corporations and states. But does anyone really think that considerations of determinism might undermine the accountability of organizations and states? Might metaphysics render the oil companies invulnerable to justifiable sanctions? If there is (as I believe) a philosophical difficulty about understanding how to think about and regard ourselves in light of the possibility of determinism, then, perhaps it is because it threatens not accountability, but rather the idea that we are criticizable in a way that goes beyond the Just Sayin’ Model. But this turns out to have relatively little to do with the ability to do otherwise. The problem lies rather elsewhere.

It is regrettable to end my paper with such large and sketchy speculations. This you might think is where the real philosophy begins, making you wish that I had gotten to these points more quickly, leaving myself some time to develop them. Since I am not literally accountable to you for writing a more brilliant and edifying paper, you cannot justifiably blame
me for this. But this is small comfort, since you can rightly criticize me, which, from my point of view, is both better and worse.